基于日志分析的数据资源授权策略评估①

汤庆新[1], 王燕妮[2], 姜宏伟[1], 张 电[1]

[1](72671 部队, 合肥 230039)
[2](国防信息学院 信息管理中心, 武汉 430000)

摘 要: 高效可靠的授权策略是实现对数据资源有效管控的关键, 但目前如何对其效率和可靠性进行测试评估尚无系统研究. 首先, 提出了一种量化开销计算方法(QCC, Quantitative Cost Calculation Method), 通过区分错误授权判定和错误拒绝授权判定的方式改进了传统的开销评估方法, 进而提出了一种以量化开销为决策依据的策略评估机制, 给出了该机制运行流程的形式化描述. 最后, 通过对某综合网管系统日志数据的分析评估, 验证了该机制的有效性.

关键词: 数据资源管理; 授权策略;量化开销; 策略评估

Data Resource Authorizaton Policy Evaluation Based on Log Data

TANG Qing-Xin[1], WANG Yan-Ni[2], JIANG Hong-Wei[1], ZHANG Dian[1]

[1](72671 PLA Troops, Jinan 250000, China)
[2](Information Management Center, Academy of National Defense Information, Wuhan 430000, China)

Abstract: An efficient and reliable data resource authorization policy is the key of effective management, and there is no systemic study on how to make an evaluation on its efficiency and reliability. Firstly, QCC (Quantitative Cost Calculation Method) Method is proposed. Compared to the current methods, QCC makes a distinction between the wrong authorization judgment and wrong refused authorization judgment. And then a policy evaluation mechanism based on QCC is put forward. At last, by carrying out real policy evaluation analysis, we verify the effectiveness of the mechanism in improving the work efficiency with the real log data of an integrated network management system.

Key words: data resource management; authorization strategy; quantitative cost; strategy evaluation

1 系统概述

各类数据资源授权和分发控制的安全性和高效性在很大程度上取决于其授权策略的选择. 目前使用较为广泛的资源授权策略主要有账号口令匹配、数字证书验证、基于短信等实现的双向验证等, 其主要原理是通过前置的认证手段阻止非法接入进而阻断非法操作. 另外, 往往还通过定期对各类日志数据的审计实现对非法授权操作的追责[1,2].

通过设置较多的验证因素(如基于短信等的双向验证)和较为严格的验证规则(如设置较短的合法验证时间窗口), 往往可以降低发生非法操作发生的可能性, 但在较为复杂的业务流程中可能因网络延迟等不可控因素引发验证超时等事件, 进而对某些时效性要求较高的业务产生严重影响[3-5]. 较为宽松的认证授权策略往往可提高运行效率, 更适用于某些时效性要求高的场合(延迟授权或拒绝授权可能带来严重后果)[1,6], 但此类策略的可靠性相对较低. 目前对在实际工作中如何评估和选择认证策略的研究较少, 尤其是对在什么条件下应该调整认证授权策略, 缺少系统全面的评估机制供系统管理人员使用.

2 应用场景分析

区分认证授权策略的运行效率和可靠性进行分析, 对于开放式、大规模的数据资源授权应用场景, 尤其

① 基金项目:第 52 期中国博士后科学基金(2012M521838)

收稿时间:2015-10-21;收到修改稿时间:2015-12-15 [doi: 10.15888/j.cnki.csa.005214]

是涉及个人隐私或资金转移的应用如网银、支付宝等, 一般应采用较为严格的授权策略, 采用如用户口令/USBkey 双因素认证等策略, 以避免大量用户带来的有意或无意违规操作[7-9]. 在较小范围内使用、对时效性要求较高的应用场景, 如军事领域的指挥控制系统, 一般应采用较为宽松的验证策略, 同时辅以高效有力的事后审计和责任追究保证策略的可靠性[10].

目前对如何选择数据资源授权策略并对其效率和可靠性进行量化评估尚无系统研究, 尤其是在何种条件下应放松授权策略以保证系统的高效运作研究较少. 一般的, 我们认为具备以下条件的应用场景(如军事领域的指挥控制系统、交通控制信号系统)中, 应用较为宽松的授权策略并加以高效的事后审计可能带来更大收益[11]:

1)业务流程复杂, 并且在业务处理中经常出现紧急且不可预知的情况, 进而导致资源授权的范围和权限难以确定, 如综合网管系统对故障源的分析和确认过程往往涉及多个设备的状态获取和日志访问权限等;

2)错误的拒绝授权可能造成严重的后果, 如在军事信息系统可能导致战斗失利, 医疗系统中可能危及病人的生命安全;

3)操作维护人员数量有限且经过严密培训, 无意误操作的可能性小;

4)受控严密, 无关人员误操作可能性小;

5)对违法操作的追溯能力强, 一般基于日志记录、视频监控等手段实现;

6)违规操作成本高, 可有效吓阻大部分恶意操作, 如交通控制系统中对违反交通信号行驶的处罚.

3 量化开销计算方法QCC

在分析了不同数据资源授权策略的应用场景后, 本节将对策略评估的依据, 即量化的授权开销进行分析研究.

3.1 相关研究

文献[12-15]介绍了数种开销量化计算方法, 但上述方法的量化粒度都较为粗糙, 主要表现在对错误判定的开销计算没有区分错误授权和错误授权判定的不同, 而在大部分情况下两者的开销并不相同, 这就导致无法对其进行对比和分析, 而授权策略的量化评估在很大程度上由错误授权和错误拒绝的不同来决定, 这就意味着现有的开销计算方法难以满足本文对策略选择评估的要求.

3.2 概念定义和前提假设

对 QCC 方法涉及的概念及表示描述如下:

1)开销(Cost): 开销整体上分为两类, 判定过程或审计过程开销(PC, Process Cost)和判定结果(RC, Result Cost)开销. 其中判定开销指判定所消耗的时间, 判定结果开销指错误或正确授权带来的后果;

2)概率(Probability): 指四种可能的授权判定结果出现的概率, 即正确真判定概率(Pct, Probability of Correct True Judgment)、正确假判定概率(Pcf, Probability of Correct False Judgment)、错误真判定概率(Pwt, Probability of Wrong True Judgment)和错误假判定概率(Pwf, Probability of Wrong True Judgment). 其中 Pct 指实际应该授权且判定结果也为授权的事件概率, Pwt 指实际应该拒绝授权但判定结果为授权的事件概率, Pcf 指实际应该拒绝授权且判定结果也为拒绝授权的事件概率, Pwf 指实际应该授权但判定结果为拒绝授权的事件概率, Pwt 和 Pcf 之和表示了可能的违规访问概率.

QCC 方法的运行基于以下两条设定:

1)QCC 方法本身可靠性为 100%, 即 QCC 运行结束后必定给出确认结果: 授权或拒绝授权, 不考虑在真实环境下可能的异常结束. 考虑到大部分错误控制机制将异常结束按照拒绝授权处理, 该假设有实际的合理性;

2)正确授权的后果开销为 0;

3)错误授权会带来判定结果开销, 且对同一资源的 wt 和 wf 开销不同.

3.3 运行流程

对不同的数据资源授权策略应用 QCC 方法, 其运行流程描述如下:

1)判定过程开销计算

① 判定步骤 *Step* 划分, 为简化计算过程, 认为每个判定步骤的开销一致;

② 第 i 次判定的开销计算: $PC_i=Count(Step)$;

③ 归一化: $\overline{PC}=\sum_{i=1}^{i=N} PC_i / N$.

2)判定后果开销计算

① 开销计算, 对于业务环节 i 有:

$$RC_i = f_i(ct)*Pct + f_i(cf)*Pcf + f_i(wt)*Pwt + f_i(wf)*Pwf$$,

其中为 f_i 结果量化函数, 根据 3.2 节关于结果开销的假

设, 有 $f_i(ct)=f_i(cf)=0$, 且区分错误真判定 *wt* 和错误假判定 *wf* 进行计算, 即:

$$RC_i = f_i(wt) * Pwt + f_i(wf) * Pwf$$

② 归一化: $\overline{RC} = \sum_{i=1}^{i=N} RC_i * P_i$, 其中 P_i 为业务环节 *i* 出现的概率.

4 基于量化开销的策略评估机制

在使用 QCC 方法完成开销量化评估的基础上, 本节从数据准备和运行流程两个方面具体描述了策略评估机制的使用方法.

4.1 数据准备

数据准备主要包括两个方面的内容, 一是业务流程环节的形式化描述, 二是对日志数据的统计分析.

4.1.1 业务流程环节的形式化描述

1)某项包含 n 个环节的业务 $Service_i$:

$Service_i \subset \{Step_1, Step_2, \ldots, Step_N\}$

2)环节 *j* 所涉及的资源:

$Resource(Step_j)=\{RS_1, RS_2, \ldots, RS_M\}$

3)错误资源授权开销

$$Cost(RS_i)=\begin{cases}\alpha_i, \text{错误授权开销}\\ \theta_i, \text{错误拒绝授权开销}\end{cases}$$

4.1.2 日志数据的统计分析

通过对资源授权日志进行统计分析, 根据 3.3 节给出的计算公式, 获取判定过程的判定开销(*Step* 的数量), 完成各业务环节出现的概率、*Pwt* 等概率值的统计计算.

4.2 运行流程

策略评估机制的具体运行流程描述如下:

1)分析业务流程, 完成业务流程形式化描述;

2)分析获取每个 *Step* 涉及的需要进行授权判定的资源;

3)完成各 *Step* 出现概率及各 *Step* 涉及资源的开销计算函数统计评估;

4)根据 1)和 2)的数据, 完成结果量化函数 *f* 的实际映射, 其中:

$$f_i(wt) = \sum RS_i * Cost(RS_i) = \sum RS_i * \alpha_i,$$

其中 $RS_i \in Resource(Step_i)$

$$f_i(wf) = \sum RS_i * Cost(RS_i) = \sum RS_i * \theta_i,$$

其中 $RS_i \in Resource(Step_i)$

5)完成开销数据的计算和归一化, 对事前预防型策略完成判定过程开销和判定结果开销的综合;

6)根据量化开销数据, 完成资源授权策略的选择.

5 实验和数据分析

结合某电信运营商综合网管系统近 1 年的系统操作维护日志, 验证本文所述资源授权策略选择机制的有效性.

综合网络系统的业务概要上可分为两类: 一是对网络系统中各设备状态的查询和监控, 二是对各设备的配置维护. 其中, 状态查询和监控操作的绝大部分由软件自动执行, 而配置维护操作需要人工执行. 显然, 查询和监控操作发生错误并不会直接影响业务运行, 相对的, 设备的配置维护操作若发生错误授权, 带来的影响可能较大. 分别对查询监控和配置维护相关日志数据进行统计分析, 得到表 1 和表 2.

表 1 查询监控日志统计数据

序号	项目	数据	说明
1	条目总数	13114734	大量自动查询操作
2	授权数	13112431	
3	拒绝授权数	2303	
4	ct	13112405	
5	wt	26	账号权限错误引发
6	cf	17	误操作引发
7	wf	2286	数据库超时等引发

表 2 配置维护日志统计数据

序号	项目	数据	说明
1	条目总数	247	系统升级扩容等
2	涉及资源	设备配置文件	
3	授权数	243	
4	拒绝授权数	4	
5	ct	243	
6	wt	0	
7	cf	4	误操作引发
8	wf	0	

由表 1 和表 2 的统计数据可以看出, 综合网管系统发生的 *cf* 均由内部人员误操作引发, 没有发现外部人员非法操作且数量占比少. 查询监控日志中存在少量 *wt*, 主要原因为人员岗位变动但账号授权范围变更延迟. 查询监控日志中同时存在一定数量(但占日志总体的比例仍然很小, 为 0.0174%)的 *wf*, 主要由其本身运行或认证数据库超时造成.

下面分析和设置各类违规操作的开销值(10 分制). 查询监控业务涉及的数据资源主要是各网络设备的运行状态, 由表 1 可知, 对其的非法访问(*wt* 和 *cf*)均为内部人员误操作导致, 不会对系统自身和业务造成影响, 但有一定的信息泄露风险, 开销设置为 1. 但对设备状

态访问的错误拒绝 *wf* 会导致设备告警, 进而引发告警工单、工单处理等事件, 会造成一定的人力开销并有一定可能因误判造成进一步损失, 但也不会对系统自身和业务造成影响, 开销设置为 2.

配置维护业务主要涉及各网络设备的配置文件, 由表 2 数据可知, 设备配置维护操作的数量少, 没有 *wt* 和 *wf* 的发生, 仅有的 4 次 *cf* 均由键盘输入错误引发. 由于网络设备的配置维护操作中认证授权所占时间比例一般较小, 错误拒绝一般只是引发再次认证, 不会对业务流程造成影响, 开销设置为 0, 而错误授权可能导致设备配置文件修改, 根据设备类型不同, 可能引发不同的后果, 对于旁路设备, 仅故障不会对其他业务造成影响, 开销设置为 4, 对于串联设备, 会引发业务中断, 若没有相应的备份冗余设备, 可能引发全网故障, 开销设置为 10.

分别对两类业务应用 QCC 方法, 其数据如表 3 所示. 由表可知, 根据量化的开销对比, 对于查询监控操作(操作数量大但每次影响较小), 应采取较为宽松的授权策略, 取消内置在程序中或显示输入的认证操作(操作数量大但每次影响较大)从而尽可能提高授权策略的运行效率, 而关系到业务能否正常运行的配置维护操作, 采取双因素认证、双向认证等较为严格的授权策略, 从而尽可能提高策略的可靠性.

表 3 QCC 开销数据

项目	开销	
	宽松策略	严格策略
查询监控	$3.5*e^{-4}$	$3.3*e^{-6}$
配置维护	0	0.45

6 结束语

本文针对现有数据资源授权策略各自的优点和不足, 提出了一种量化开销计算方法, 通过区分错误授权判定和错误拒绝授权判定的方式改进了传统的开销评估方法, 进而提出了一种以量化开销为决策依据的策略评估机制. 最后结合某综合网管系统的实际日志数据, 验证了该策略评估机制在科学策略选择以合理平衡不同策略的工作效率和可靠性的作用.

参考文献

1 Weitzner DJ, Abelson H, Berners-Lee T, Feigenbaum J, Hendler J. Information accountability. Communications of the ACM, 2008, 51(6): 82−87.

2 Sweeney L. K-anonymity: A model for protecting privacy. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, October 2002, 10: 557−570.

3 Abendroth J, Jensen CD. Partial outsourcing: A new paradigm for access control. Proc. of the Eighth ACM Symposium on Access Control Models and Technologies. 2003. 134−141.

4 Cohen E. Models for coalition-based access control. 7th ACM Symposium on Access Control Models and Technologies. 2002. 97−106.

5 Park RSSJ. Towards usage control models: Beyond traditional access control. 7th ACM Symposium on Access Control Models and Technologies. 2002. 57−64.

6 Weeks S. Understanding trust management systems. Proc IEEE Symposium on Security & Privacy, 2001: 94-105.

7 Wadlow T. Who must you trust? Queue-Security. May 2014, 12(5): 30−36.

8 伍华凤,戴新发,陈鹏.一种层次化移动 IP 接入认证机制.计算机工程,2008,24:131−133.

9 Banga M, Hsiao M. A novel sustained vector technique for the detection of hardware Trojans. Proc. 22nd International Conference on VLSI Design. IEEE CS Press. 2009. 327−332.

10 Cheng P, Rohatgi P, Keser C, Karger P, Wagner G, Reninger A. Fuzzy multi-level security: An experiment on quantified risk-adaptive access control. Proc. of the IEEE Symposium on Security and Privacy. 2007. 222−230.

11 Robertson K. Kaiser fined for employees checking medical records of octuplets. Sacramento Business Journal. 2009. 7−16,.

12 Dempsey K, Witte G, Rike D. Security and privacy controls for federal information system and organizations. Technical Report Special Publication -800-53, Revision 4. Washington, DC. 2014.

13 Yuan E, Esfahani N, Malek S. A systematic survey of self-protecting software systems. Trans. on Autonomous and Adaptive Systems, 2014, 8(4): 45−52.

14 Ni Q, Bertino E, Lobo J. Risk-based access control systems built on fuzzy inferences. Proc. of the 5th ACM Symposium on Information, Computer and Communications Security. April 2010. 250−260.

15 Benrejeb M, Sakly A, Othman KB, Borne P. Choice of conjunctive operator of TSK fuzzy systems and stability domain study. Mathematics and Computers in Simulation, 2008, 76(6): 410−421.